The Gsp-1 genes encode the 1 wheat arabinogalactan peptide

Journal of Cereal Science 74 (2017) 155–164

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

The *Gsp-1* genes encode the wheat arabinogalactan peptide

Mark D. Wilkinson [a, *], Paola Tosi [b], Alison Lovegrove [a], Delia I. Corol [a], Jane L. Ward [a], Richard Palmer [a], Stephen Powers [a], Diana Passmore [a], Gina Webster [a], Susan E. Marcus [c], J. Paul Knox [c], Peter R. Shewry [a, b, **]

[a] Plant Biology and Crop Science Department, Rothamsted Research, Harpenden AL5 2JQ, UK
[b] School of Agriculture Policy and Development, University of Reading, Reading RG6 6AH, UK
[c] Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Article history:
Received 25 January 2017
Accepted 14 February 2017
Available online 16 February 2017

Keywords:
Wheat
Arabinogalactan peptide
Grain softness protein
Immunocytochemistry

ABSTRACT

Western blotting, ELISA and ^{1}H-NMR spectroscopy showed that RNAi down-regulation of the wheat *Gsp-1* gene resulted in reduced contents of both arabinogalactan peptide (AGP) and grain softness protein (*GSP-1*) in mature wheat grains confirming that these components are encoded by the same gene. A small increase in grain hardness and decrease in the viscosity of aqueous extracts of the transgenic lines also indicated small effects on functional properties. Immunolocalisation using a novel wheat AGP monoclonal antibody in conjunction with confocal microscopy showed that the major form of AGP which was eliminated in knockout lines is located within the cell, probably in the vacuole, and not in the plasma membrane or cell wall. However, clear localisation of the AGP epitope to the plasma membrane was observed in both control and transgenic lines and probably resulted from the presence of one or more separate forms of arabinogalactan protein. The existence of such additional form(s) was also indicated by ^{1}H-NMR spectroscopy which showed that the ratio of arabinose to galactose differed between the control and transgenic lines.

1. Introduction

The *Gsp-1* genes of wheat are proposed to encode precursor proteins that are post-translationally processed to give two components: the arabinogalactan peptide (comprising 15 amino acids) and the grain softness protein (*GSP-1*) (the major form comprising 113 amino acids with 10 cysteine residues) (Elmorjani et al., 2013). However, it has not been directly demonstrated that the AGP is derived from *Gsp-1* genes. *Gsp-1* genes appear to be ubiquitous in grasses, being identified in 65 species from the five major grass subfamilies, with over 20 different *Gsp-1* alleles being reported (Wilkinson et al., 2013). The GSP protein belongs to a large family of seeds storage proteins called the "prolamin superfamily" (Shewry and Halford, 2002) and has been predicted to have a similar structure to other members of this family, including the 2S storage albumins which occur in dicotyledonous seeds and the puroindolines (Pins) of cereal seeds (Elmorjani et al., 2013).

The *Gsp-1* genes are present at the *Ha* (Hardness) locus on chromosome 5D (Chantret et al., 2005), which also comprises two genes encoding the puroindolines Pin a and Pin b. Allelic variation in the expression and/or sequences of the Pin proteins has been shown to account for 60–80% of the variation in endosperm texture (hardness) in bread wheat (*T. aestivum* L.) cultivars (Turner et al., 2004). Grain hardness has a major impact on the processing properties of bread wheat with hard texture being preferred for bread making and soft texture for biscuits, cakes and pastries (Morris and Rose, 1996). Although Pins have been studied in considerable detail, little is known of their structures and biological roles, nor about the mechanisms that determine grain texture. Although GSP was initially reported to affect grain texture (Jolly et al., 1996), this has since been suggested to result from the tight linkage of the *Gsp-1* and *Pin* genes with no direct effect of GSP itself (Tranquilli et al., 2002).

Fincher and Stone (1974) showed that the wheat arabinogalactan peptide (AGP) has a mass of about 22,000 and consists of 80% polysaccharide. More recent studies have shown that wheat AGP comprises about 0.4% of the dry weight of the wheat endosperm

* Corresponding author.
** Corresponding author. Plant Biology and Crop Science Department, Rothamsted Research, Harpenden AL5 2JQ, UK.
E-mail addresses: mark.wilkinson@rothamsted.ac.uk (M.D. Wilkinson), peter.shewry@rothamsted.ac.uk (P.R. Shewry).

http://dx.doi.org/10.1016/j.jcs.2017.02.006

(Loosveld et al., 1997) with similar contents of AGPs in related cereals, namely barley (0.28%), rye (0.21%), durum wheat (0.28%) and triticale (0.32%) (Van den Bulck et al., 2005). Van den Bulck et al. (2005) showed that the peptide moiety of AGP was identical in sequence to the first 15 amino acids predicted for the protein precursor of *GSP-1* (allowing for post-translational hydroxylation and glycosylation of the 3 proline residues) (Van den Bulck et al., 2005). The cleavage between the AGP and *GSP-1* sequences is at the C-terminal side of an aspartic acid residue, which is consistent with the action of a legumin-type cysteine endopeptidase similar to those responsible for the post-translational processing of storage globulins and 2S albumins in the vacuoles of dicotyledonous seeds. It can therefore be hypothesised that the AGP/GSP precursor protein is synthesised on the endoplasmic reticulum (ER), transferred into the lumen (a signal sequence of 19 amino acids also being predicted) and post-translationally processed (proline hydroxylation, O-glycosylation) in the ER and Golgi before being proteolytically cleaved in the vacuole (Wilkinson et al., 2013). In support of this hypothesis, Elmorjani et al. (2013) showed that the GSP precursor protein is further proteolytically cleaved at the *N*- and *C*-termini to give four forms comprising between 110 and 113 residues.

Although substantial amounts of both AGP and GSP are present in the starchy endosperm (white flour) of the mature grain nothing is known about their locations in the cells, or their biological functions, with no evidence for the widely held assumption that AGP is located in the cell wall. Furthermore, nothing is known about their effect on grain processing or human health, although AGP is present in similar amounts to water-soluble fibre components.

We have therefore generated and analysed transgenic wheat lines with RNAi-mediated suppression of *Gsp-1* transcripts. ^{1}H-NMR analysis, ELISA and western blotting showed decreases in the contents of both AGP and GSP, demonstrating for the first time that these two components are indeed encoded by the same genes. Immunolabelling shows that the major form of AGP encoded by *Gsp-1* has an intracellular location, probably within the vacuoles that give rise to protein bodies, while analysis of mature grain shows small but statistically significant effects on both grain texture and on the viscosity of aqueous extracts of endosperm flour.

2. Materials and methods

2.1. RNAi construct preparation, transformation and confirmation of transgenic wheat lines

An RNAi construct to down-regulate *Gsp-1* under the control of the starchy endosperm-specific *HMW1Dx5* promoter was created using a *Bgl*II/*BamH*1 cloning strategy as described by Nemeth et al. (2010), using 344-bp fragments (+1 to +344bp of the sequence) based on a cDNA sequence from *T. aestivum* cv. Cadenza AGP1; (Wilkinson et al., 2013). This fragment was chosen as it corresponds to the most highly expressed *Gsp-1* sequence. The fragment was generated by PCR (Phusion *Taq* polymerase from Thermo Scientific, Fisher Scientific - UK Ltd Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG United Kingdom), using the following PCR primers GSPRNAiF and GSPRNAiR (Supplementary Table S1). The full nucleotide and protein sequences of the *Gsp-1* sequence are shown in Supplementary Figs. S1a and S1b respectively.

Sequencing of the constructs was carried out using the BigDye Terminator Version 3.1 Cycle Sequencing Kit (Applied Biosystems, Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK), with construct specific primers i.e. M13F, Rab1 and adh3R (Supplementary Table S1) used to confirm correct orientation of RNAi fragments. All reactions were analysed at Source Bioscience (Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK).

Wheat transformation was carried out by particle bombardment (PDS1000; Bio-Rad) of immature scutella (10–14 days after anthesis) of cv. Cadenza, co-bombarded with the pAHC20 plasmid, containing the selectable marker gene *bar* driven by the constitutive ubiquitin promoter from maize, as described by Sparks and Jones (2009).

Genomic DNA was extracted from leaf material using a Promega Wizard kit (Promega (UK) Ltd, Southampton, Hampshire, SO16 7NS, UK). PCR analysis to confirm the presence of the transgenes was carried out using the Rab1 and GSPRNAiR primers described above for the RNAi wheat lines (Supplementary Table S1). The reactions were performed in 25 μl using a 1.1× ReddyMix™PCR Master Mix (1.5 mM $MgCl_2$) from Thermo Scientific (ABgene House, Blenheim Road, Epsom, Surrey, KT19 9AP, UK), also containing ~200 ng of genomic DNA and 0.8 μM of each primer. The cycling conditions were 96 °C for 5 min followed by 32 cycles of 96 °C for 30 s; 56 °C for 30 s; 72 °C for 1 min 30 s and the extension of 72 °C for 10 min for PCR reactions. PCR products were analysed on 1.0% (w/v) agarose gels, stained with ethidium bromide and visualised by UV light.

Zygosity analysis of T1 plant material was carried out as described by Nemeth et al. (2010) to identify transgenic lines with segregation patterns consistent with a single insertion locus.

2.2. Plant growth

Homozygous and azygous (null) T2 segregants descended from the same original RNAi transformants were grown in four replicate pots, with four plants per pot, in a four block design (one replicate of each line i.e. transgenic and null per block) in temperature controlled GM glasshouse rooms with 18 °C–20 °C day and 14 °C–16 °C night temperatures and a 16-h photoperiod provided by natural light supplemented with banks of Son-T 400 W sodium lamps (Osram, Ltd.) giving 400 to 1000 μmol m^{-2} s^{-1} photosynthetically active radiation.

2.3. RNA extraction, quantitative reverse transcription-PCR analysis and semi-quantitative reverse transcription-PCR analyses of RNAi wheat lines

Tissue samples enriched in starchy endosperm cells were isolated from T3 developing grain of wheat (*T. aestivum* cv. Cadenza) RNAi lines at 14 d.p.a by gentle squeezing to remove the pericarp. RNA was extracted as reported by Wilkinson et al. (2013) and DNase (Promega (UK) Ltd, Southampton, Hampshire, SO16 7NS UK) used to remove contaminating DNA. First-strand cDNA synthesis was carried out with a Superscript III Reverse Transcriptase Kit (Invitrogen, Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK). The cDNA synthesis step was carried out as follows in a final 20 μl volume: 2 μg equivalent of DNase treated RNA adjusted to a 12 μl volume with sterile distilled water, 1 μl of 100 μM Oligo-dT primer, 1 μl of 10 mm dNTPs; 65 °C for 5 min followed by 1 min incubation on ice. This was followed by the addition of 4 μl of a 5× first strand cDNA synthesis buffer, 1 μl of 0.1M DTT and 1 μl of Superscript III Reverse Transcriptase; 50 °C for 60 min and 70 °C for 15 min.

Down regulation of the transcript was measured using the SYBR Green Jumpstart™ Taq Readymix™ for quantitative reverse transcription-PCR (Sigma-Aldrich Company Ltd. Dorset, UK). The reaction was carried out in a 20 μl mixture containing 1× Green Jumpstart™ Taq Readymix™, 5 μl 0.4 μM of each primer with the ROX reference dye provided, 5 μl cDNA working volume (1:15 dilution of original cDNA reactions). The following temperature profile was used: an initial denaturation step of 95 °C for 10 min,

followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min with an added dissociation stage of 95 °C 15 s; 60 °C for 1 min; 95 °C for 15 s and 60 °C for 15 s. An annealing temperature of 59 °C was used for the *Pina* and *Pinb* experiments. All reactions were carried out in 96 well plates and analysed using the Applied Biosystems 7500 machine and software v2.0.5. Down regulation in the *Gsp-1* RNAi lines was determined using the primers GSPqPCR3 and GSPqPCR4, resulting in an amplicon of 77bp. *Pina* and *Pinb* transcripts were determined using PinaqPCR4F + PinaqPCR4R and PinbqPCR4F + PinbqPCR4R, resulting in an amplicon of 139bp and 135bp respectively (Supplementary Table S1).

Three reference genes were used to normalize expression: *Ta*2526 (Prosm), a stably expressed EST from grain (primers prTYW19 and TYW20), glyceraldehyde-3-phosphate dehydrogenase (primers prTYW422 and prTYW423) (*Ta*GAPDH), and succinate dehydrogenase (primers prTYW424 and prTYW425) (*Ta*SDH). All primer sequences are given in Supplemental Table S1.

Following qRT-PCR, normalised relative quantity (NRQ) data were calculated for the target genes of interest by using the formula: $NRQ = E_T^{-ctT}/(E_{R1}^{-ctR1}.E_{R2}^{-ctR2}.E_{R3}^{-ctR3})$; where E_T, E_{R1}, E_{R2}, E_{R3} are the efficiencies of the target (*Pina*, *Pinb* or *Gsp-1*) and the three reference genes used (Prosm, *Ta*SDH and *Ta*GAPDH); *ctT*, *ctR1*, *ctR2*, *ctR3* are the corresponding numbers of cycles at threshold fluorescence set for the reactions; and the denominator of the expression is the geometric average of the relative expression of the three reference genes. As these data have heterogeneity of variance, a transformation, $\log_2(1/NRQ)$ (Rieu and Powers, 2008), was applied prior to analysis. Linear mixed modelling was applied using the method of residual maximum likelihood (REML) as implemented in the GenStat (17th edition, VSN International Ltd, Hemel Hempstead, UK) statistical package to test for the overall effect of the transgene and then to test differences between lines (F-tests) for each target gene. The analysis took account of the randomised block design used for growing the plants and the plates used at the laboratory stage of the experiment. Means were compared using the standard error of the difference (SED) on the residual degrees of freedom from the ANOVA thus invoking the least significant difference at the 5% level of significance.

Semi-quantitative reverse transcription-PCR cDNA synthesis was carried out as described above and full length *Pina* and *Pinb* products were screened using the primers shown in Supplementary Table S1 and PCR conditions as described above.

2.4. Preparation of flour for biochemical analysis

Grain were conditioned to 15.5% water and milled using a Micro Scale Labmill FQC-2000 (METEFÉM SZÖVETKEZET, Hungary, 1047 Budapest, TINÓDI U. 28-30.). Analyses were carried out on the flour fraction below 150 μM.

2.5. Generation of monoclonal antibodies to wheat AGP and ELISA

AGP was isolated from Cadenza white flour as described in Tryfona et al. (2010), freeze-dried and stored at −80 °C. Immunization of rats, preparation of hybridoma cell lines and cell cloning were performed as described in Willats et al. (1998). Two male Wistar rats were injected with ~100 μg each of wheat AGP in complete Freund's adjuvant administered sub-cutaneously on day 0, with the same amount administered with incomplete Freund's adjuvant on days 35, 70 and 98. On day 147, a rat was given a pre-fusion boost of 100 μg immunogen in 1 ml PBS by intraperitoneal injection. Spleen lymphocytes were isolated three days later and fused with rat myeloma cell line IR983F (Bazin, 1982). Hybridoma cell lines were selected by ELISA using the wheat AGP as antigen and this led to the isolation of 7 rat monoclonal antibodies that bound to the wheat AGP that were all of the IgM class. An antibody-, designated LM30 was selected as the most highly specific for wheat AGP (Supplementary Fig. S2).

For ELISA, 1 mg flour was incubated in 1 ml PBS (phosphate-buffered saline) overnight at RT with rocking. The supernatant was then incubated overnight in Nunc Immunosorb microtitre plate wells (100 μl/well) at equivalents to the original 1 mg/ml and 5-fold dilutions to 6.4 ng/ml. Microtitre plate wells were blocked with milk protein in PBS (MP/PBS) to prevent non-specific binding and then incubated with LM30 (10-fold dilution in MP/PBS) for 2 h. After washing the plates were probed with anti-Rat-IgG-peroxidase and developed with a chloronaphthol-based substrate and absorbance (450 nm) determined. ELISA was performed in sixteen 96-well microtitre plates, each plate accounting for all the observations for four samples, and each original block (of plants) being represented by four plates. Statistically, the data for one sample (from 24 wells on a given plate) were all non-independent and, to account for this, a two-stage method of analysis was used (Mead et al., 2003). As stage one, the absorbance data for each sample were modelled over the range of dilutions on the log (to base *e*) scale, using a logistic model:

$$y = \frac{C}{(1 + \exp[B(\log_e(dilution) - M)])}$$

Where *C* is the asymptotic absorbance, *B* is the effective rate of increase and *M* is the dilution providing half of the asymptotic absorbance. This curve was chosen as the best descriptor for the majority of samples, in comparison to an asymmetric Gompertz, or a critical exponential, model. Stage two analysed the sets of parameter estimates, *C*, *B* and *M* from the fit of the logistic model across the treatment structure of the experiment to extract the significance of differences between genetic types and lines within types. For this, the method of residual maximum likelihood (REML) was used to fit a linear mixed model to the sets of estimated parameters, accounting for blocks and plates effectively nested within blocks as random terms (variance components) and testing (F-tests) the main effect of types and lines nested within types as fixed terms. This approach was used because the observations of lines were unbalanced across the plates, there being four lines per plate. The relevant predicted means of the estimated parameters were output along with standard error of the difference (SED) values. The predicted mean logistic curves for the lines were then drawn, using the predicted means for the three parameters. The GenStat (14th edition, © VSN International, Hemel Hempstead, UK) statistical package was used for the analysis.

2.6. ^{1}H-NMR analysis of endosperm flour

Endosperm flour (30 mg) was extracted at 50 °C in D_2O:CD_3OD (80:20, 1 ml) containing d_4-TSP (0.01% w/v) as internal standard (Baker et al., 2006). ^{1}H-NMR data was collected at 300 K on a Bruker Avance (Coventry, UK) spectrometer operating at 600.05 MHz using a Selective Inverse 5 mm probe. Spectra were collected using 128 scans of 64,000 data points with 7184 Hz spectral width. A water suppression pulse sequence with a 5 s relaxation delay was used to suppress residual the HOD signal. Spectra were Fourier-transformed using an exponential window with 0.5 Hz line broadening and were automatically phased and baseline corrected within the instrument software. ^{1}H chemical shifts were referenced to d_4-TSP at δ 0.0 and spectra were reduced in Amix software (Bruker Biospin) to ASCII files containing integrated regions of equal width (0.001 ppm). For AGP units, arabinose and galactose equivalents were quantified using their anomeric signals using the regions δ 5.2585–5.2244 and δ 4.5480–4.505 respectively. These

regions were confirmed *via* comparison to an authentic standard of AGP purified from wheat.

2.7. Western blot analysis

200 mg flour samples were extracted twice with 500 μl distilled water (250μl/100 mg sample) for 1 h at room temperature with shaking. Samples were diluted with NuPAGE LDS Sample Buffer (4×) and separated (30 μl) on a precast gradient gel (Bolt™ 4–12% Bis-Tris plus gels) using MES buffer (200 V, ~30 min). Gels were equilibrated in Bolt transfer buffer and then blotted on nitrocellulose membrane (Novex, 0.2 μm pore size) using a Novex Semi-Dry Blotter apparatus, following manufacturer instructions.

Efficiency and homogeneity of transfer were checked by staining with Ponceau S (Sigma-Aldrich). Western blot analysis was carried out as described in Tosi et al. (2009). Primary antibodies used were RAT monoclonal LM30, raised against wheat AGP (see above), used at a 1:500 dilution, and rabbit polyclonal -GSP (kindly provided by Dr Didier Marion, INRA, Nantes, France), raised against wheat GSP, used at 1:2000 dilution. Secondary antibodies, anti-rat AP-conjugated and anti-rabbit AP conjugated, were from Sigma-Aldrich.

2.8. Relative viscosity

The relative viscosity of aqueous extracts was determined using a AVS® 370 Viscometer, SI Analytics, Germany at 30 °C fitted with 2 ml, 0.43 mm Micro Ostwald capillary and WinVisc 3.83 software. All samples were prepared as described in Saulnier et al. (1995).

2.9. Microscopy

Caryopses were harvested from the middle third of the spikelet at 12–15 d.p.a. and prepared immediate using the Leica Microsystems EM HPM100 high pressure system. Transverse medial sections of wheat grains (less than 300 μM thick) were hand cut in a drop of 2% sucrose, 0.1 M Tris, under the microscope and cryofixed at a rate of 20,000 K per sec. Freeze substitution of cryofixed samples was performed in an EM AFS unit (Leica Microsystems) at −85 °C for 48 h in 0.5% uranyl acetate in dry acetone. The samples were then infiltrated at low temperature in LR White resin and polymerized with a UV lamp following the procedure described in Tosi et al. (2009).

Semi-thin sections of 1 μm thickness were cut using a Reichert-Jung ultramicrotome, collected in drops of distilled water on multi-well slides coated with Poly-L-lysine hydrobromide (Sigma P1399), and dried on a hot plate at 40 °C.

Slides with LR White-embedded grain sections were pre-incubated (50 μl drop/well) in 5% (w/v) milk powder (Marvel products) in 1xPBS at pH 7.0 for 60 min, then incubated for 2 h with the LM30 primary rat antibody diluted 1:5 in 1% BSA in PBS, 0.1% Tween 20 (PBST). Slides were rinsed three times for 5 min with PBST, then incubated for 2 h, in the dark, with secondary antibody (anti-rat Alexa 568 conjugated, Invitrogen) diluted 1:200 in PBS, 1% BSA. Slides were then rinsed three times with PBS, and counter-stained with 1% (w/v) Calcofluor White solution. Sections were then mounted in Citifluor AF1 glycerol based antifade mountant and analysed on a Zeiss Axiophot fluorescence microscope equipped with a Retiga Exi (Qimaging) camera.

3. Results and discussion

3.1. Quantitative PCR analysis of transgenic RNAi wheat lines

Eight transgenic lines of cv Cadenza were generated expressing a *Gsp-1* RNAi construct under the control of the starchy endosperm-specific *HMW1Dx5* gene promoter (Supplementary Table S2). Our sequence analysis of cDNA based clones (unpublished results) from *T. aestivum* cv. Cadenza, using endosperm enriched RNA, showed that *Gsp1a* was the most abundant transcript detected, at 72% of the total, with *Gsp1b* at 25% and other minor forms including *Gsp1c* at 3%. The RNAi construct designed was therefore based on the *Gsp1a* transcript (344bp of a 495bp transcript length; Supplementary Figs. S1a and S1b), which shares 94% sequence identity with

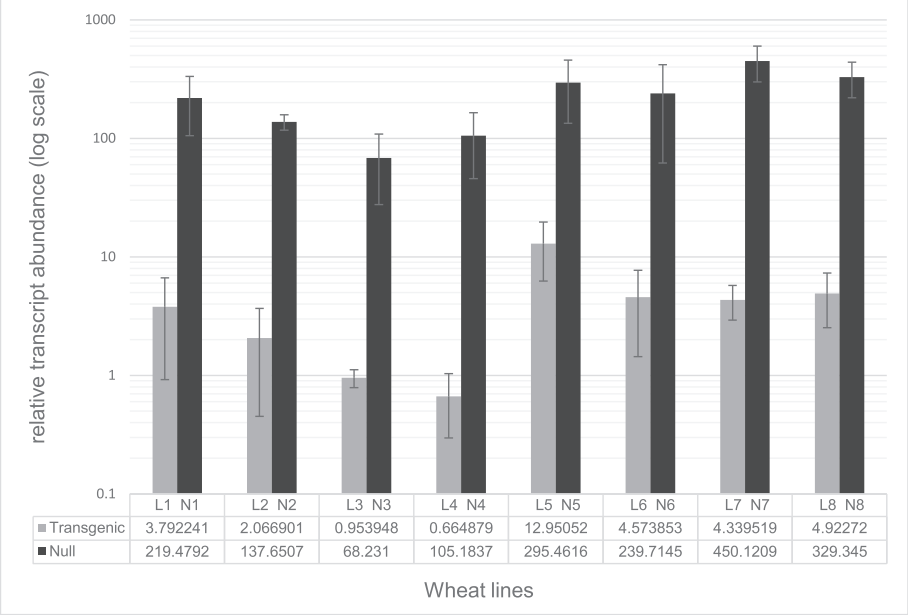

Fig. 1. Effects of RNAi transgenes on *Gsp-1* transcript abundance determined by quantitative reverse transcription-PCR from developing endosperm at 14 d.p.a. L1-L8 are transgenic lines, N1-N8 the corresponding null lines. Three biological replicates were analysed for each RNAi lines and its corresponding null. Due to large range difference in transcript abundance detected, values in the above graph are shown using a log scale. Error bars represent SE; $p < 0.05$.

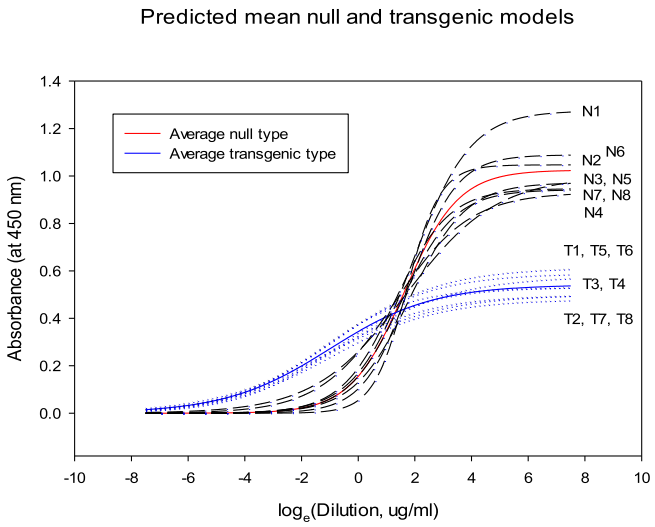

Fig. 2. Plots of predicted mean curves carried out on an ELISA analysis of LM30 binding to a range of flour dilutions for eight transgenic homozygous lines for *Gsp-1* RNAi (T) and their corresponding nulls (N) ($p < 0.001$).

Gsp1b and 99% identity with *Gsp1c*. Due to this high identity between the sequences the RNAi construct would be expected to have a profound knockdown effect on the transcript levels for *Gsp1b* and *Gsp1c* as well as *Gsp-1a* (Supplementary Fig. S1c).

The abundances of *Gsp-1* transcripts in developing endosperms (14 days after anthesis) of the transgenic and null segregants from the same transformation events were determined by quantitative reverse transcription-PCR, showing strong effects of the transgene for all lines ($p < 0.05$, LSD). The fold change reductions in the RNAi lines compared to their corresponding nulls ranged from ×52 (L6) to ×158 (L4) (Fig. 1; Supplementary Tables S3a and S3b). Line 4 had the lowest level of *Gsp-1* transcripts and was significantly different from all the other lines ($p < 0.05$, LSD) except lines 2 and 3 (Supplementary Tables S3a and S3b).

Cadenza expresses the *Pina-D1a* and *Pinb-D1c* alleles which show 42% sequence identity to the *Gsp-1* transcript sequence (Rahman et al., 1994). The abundances of *Pina* and *Pinb* transcripts in developing endosperm were, therefore, also determined by quantitative reverse transcription-PCR. Line 3 was not included in this analysis as semi-quantitative reverse transcription-PCR experiments had shown extremely low levels of *Pina* and *Pinb* transcripts (Supplementary Fig. S3). Statistical analysis of the data for the other lines showed differences in the abundances of *Pina* transcripts between the lines ($p < 0.001$, F-test) but only small differences in *Pinb* transcripts ($p = 0.065$, F-test). However, there were no statistically effects on the *Pina* or *Pinb* transcript levels when comparing the null and transgenic lines, and no effects of the transgene for individual lines (Supplementary Tables S3a–3c). Hence it was concluded that phenotypic differences observed between pairs of transgenic were not secondary effect resulting from the suppression of *Pin* gene expression.

3.2. Effects of RNAi on amounts of AGP and GSP

T2 seeds of eight pairs of homozygous and null lines for the AGP/*Gsp-1* RNAi construct were grown in a randomised block experiment providing T3 seeds for biochemical analysis.

In order to carry out quantitative analysis of AGP, a specific monoclonal antibody, designated LM30, was generated using the isolated wheat AGP as immunogen. The binding of LM30 to wheat AGP was highly sensitive to the action of arabinofuranosidases (Supplementary Fig. S2) indicating that it bound to terminal arabinose residues decorating the galactan. ELISA assays (16 lines, 4 replicate blocks) of aqueous extracts of flour fractions using the LM30 antibody showed reduced levels of AGP in the transgenic RNAi lines compared to the corresponding null controls over a range of flour dilutions. Compared to the corresponding nulls, the reductions in AGP content ranged between 38% (L5) and 54% (L2) (Fig. 2) (the full ELISA dataset and means table are given in Supplementary Tables S4a and S4b). REML revealed a main effect ($p < 0.001$, F-test) of null versus transgenic with no significant differences between lines. The dilution at which half the maximal absorbance is reached was also 14-fold lower for the transgenic lines: means of 5.135 μg/ml for the null and 0.365 μg/ml for the

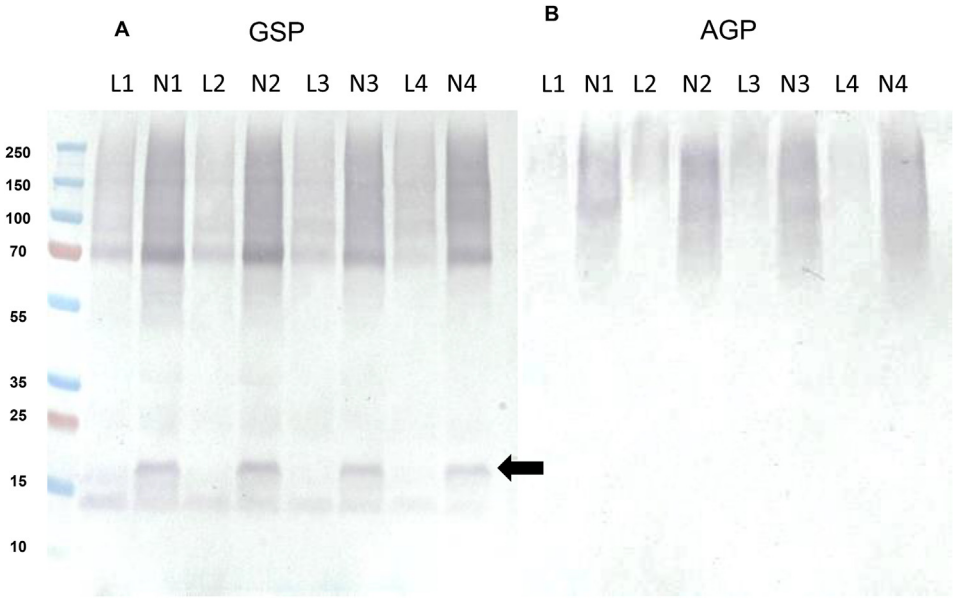

Fig. 3. Western blotting of aqueous extracts of pairs of transgenic *Gsp-1* RNAi and corresponding null lines using the polyclonal antibody to GSP (left) and the monoclonal antibody LM30 specific to the wheat arabinogalactan peptide (AGP) (right). L1-L4 are transgenic lines, N1-N4 the corresponding null lines. The GSP band is indicated by an arrow.

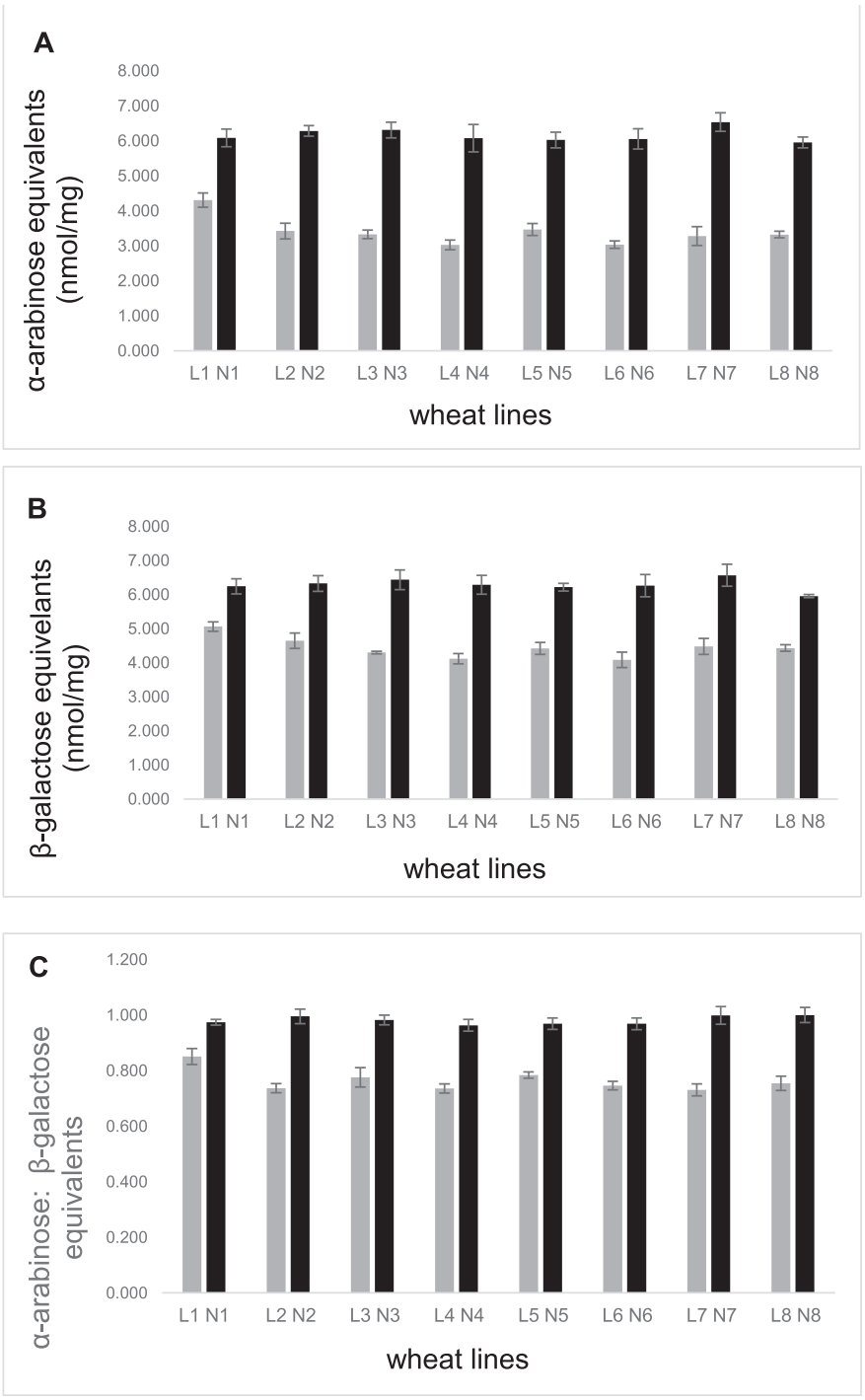

Fig. 4. Effects of the RNAi transgene on the amounts of arabinose **(A)**, galactose **(B)** and the ratios of α-Ara/β-Gal concentrations **(C)** in AGP of eight *Gsp-1* RNAi transgenic homozygous lines compared to their corresponding null controls. L1-L8 are transgenic lines, N1-N8 the corresponding null lines. Four replicate samples were used for all measurements ($p < 0.05$, LSD).

transgenic lines (calculated by back-transforming the *M* values shown in Supplementary Tables S4a and S4b). The predicted means for the estimated parameters, with standard errors, for all 16 lines are given in Supplementary Table S4c. The predicted mean logistic curves in Fig. 2 show significant differences between the transgenic and null lines, whereas the variation within these types was not statistically significant according to the REML analysis.

The reduction in AGP content in the transgenic compared with the control lines was confirmed by western blot and ^{1}H-NMR analysis. Western blotting (Fig. 3A) with the LM30 antibody revealed a number of poorly resolved bands with masses above about 70,000 (based on the mobilities of the marker proteins). The intensity of the reaction was clearly reduced in the RNAi lines, but some immunoreactive material remained, particularly in the higher molecular mass range. The presence of multiple bands is consistent with O-glycosylation, which results in a range of glycan structures, and with a previous detailed study of the glycan component of wheat AGP (Tryfona et al., 2010). However, the molecular masses

determined for glycosylated proteins by SDS-PAGE do not reflect true masses, with the wheat AGP having been reported to have masses ranging from about 22,000 to 70,000 depending on the method used (Fincher and Stone, 1974; Loosveld et al., 1997; Van den Bulck et al., 2005).

^{1}H-NMR analysis of extracts of white flour made with deuterated solvent (D_2O:CD_3OD (80:20)) showed statistically significant decreases in the integrated peak areas for chemical shifts corresponding to anomeric peaks of both arabinose and galactose units of AGP (Supplementary Tables 5a and 5b). These are quantified against an internal standard (d_4-TSP), as arabinose and galactose equivalents in Fig. 4A and B, the means being 6.29 ± 0.063 and 4.44 ± 0.111 for arabinose in the null and transgenic lines, respectively, and 6.17 ± 0.069 and 3.40 ± 0.142 for galactose. The decrease in arabinose equivalents in the RNAi lines was greater than that in galactose equivalents, resulting in a decrease in the arabinose to galactose ratio (from a mean of approx. 1 to approx. 0.75) (Fig. 4C). Our ^{1}H-NMR analysis of wheat AGP also showed an arabinose: galactose ratio of about 1, whereas monosaccharide analyses of purified AGP have shown lower ratios, of about 0.5 (Tryfona et al., 2010) and between 0.63 and 0.72 (Loosveld et al., 1997).

These analyses indicate that amount of the major form of AGP recognised by the antibody was reduced in the transgenic lines. However, it is clear that other forms of arabinogalactan proteins or peptides were also recognised by the antibody. This is apparent from the western blotting which showed the presence of poorly resolved high molecular mass components in the control and transgenic lines, and by ^{1}H-NMR spectroscopy which showed that the form(s) remaining in the transgenic lines had a lower ratio of arabinose to galactose.

The GSP component released from the precursor protein encoded by *Gsp-1* is poorly characterised at the protein level with a single published detailed study (Elmorjani et al., 2013). Elmorjani et al. (2013) raised and characterised a polyclonal antibody to their GSP preparation and western blotting of aqueous extracts of our lines with this antibody showed a clear decrease in binding to a component of mass about 18,000 in the transgenic lines compared to the corresponding control lines (see arrow in Fig. 3B). The mass of this component (based on the mobility of marker proteins) is above that reported for GSP by Elmorjani et al. (2013), who showed that the mature protein is processed to give mature forms of mass 12,639.5, 12,740.5, 12,827.5 and 12,884.5. This could result from incomplete processing, or reflect the inaccuracy of SDS-PAGE for mass determination. The same aqueous extracts were used for blotting with the LM30 AGP antibody and the GSP antibody. Comparison of the blots shows that the AGP antibody did not bind

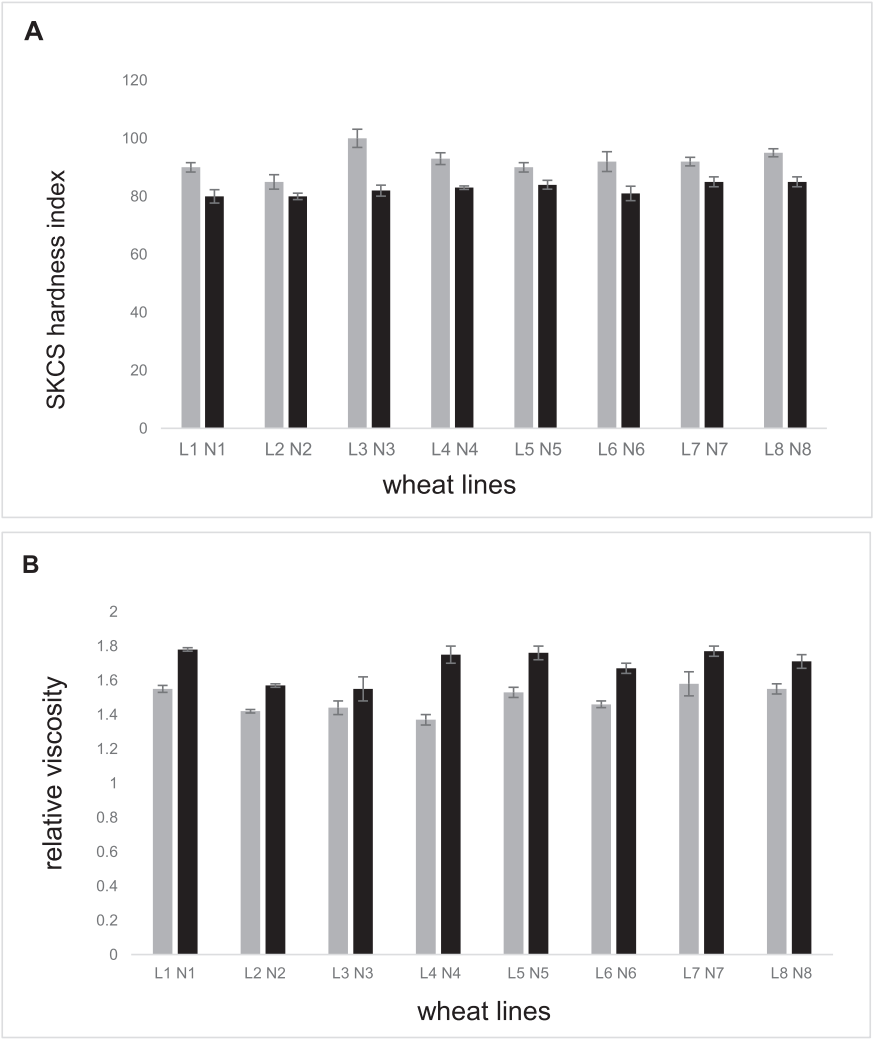

Fig. 5. Effects of the RNAi transgene on grain hardness (A) and relative viscosity (B) for the eight *Gsp-1* RNAi homozygous lines compared to their corresponding null controls. Four replicate samples were used for relative viscosity and 300 individual grains from a randomised four block experiment for grain hardness ($p < 0.05$, LSD).

to the band corresponding to GSP, or to other low molecular mass components.

3.3. Effect on grain texture

Determination of grain hardness using the Perten 4100 Single Kernel Characterization System (SKCS) showed that all transgenic lines except line 2 had harder grains ($p < 0.05$, LSD) than the corresponding null lines (Fig. 5A), the overall means being 82.46 for the null lines and 92.12 for the transgenic lines. Statistical analysis also showed an interaction between line and type ($p = 0.060$, F-test) for each line. Although the effect of RNAi suppression on hardness is modest (about 10 SKCS units), we consider that it is a genuine result of the suppression of *Gsp-1* expression, because the quantitative reverse transcription-PCR showed no effects of the suppression on the abundances of endogenous transcripts for *Pina* and *Pinb,* the major determinants of grain texture. However, we cannot conclude whether the reduced hardness relates to effects on the amounts of AGP, GSP or both components encoded by *Gsp-1*.

3.4. Relative viscosity

The viscosity of aqueous extracts is an important trait determining the end use quality of wheat, as it affects food processing, feed conversion efficiency in livestock, the quality for distilling and biofuel production, and also has impacts on human health. The major determinant of viscosity in extracts of wheat flour is the concentration and structure of water-extractable (WE) AX, and viscosity has been used as a measure of WE-AX in genetic studies (Charmet et al., 2009). However, AGP would also be expected to contribute to viscosity, as it is water-soluble and accounts for about 0.4% of the starchy endosperm dry weight (Loosveld et al., 1997).

Determination of the relative viscosity of aqueous extracts of the transgenic and null lines using a capillary viscometer showed a clear and statistically significant reduction in relative viscosity in the transgenic lines, with the mean decreasing from 1.694 to 1.494 (SED = 0.0211 on df = 45; LSD (5%) = 0.0424) (Fig. 5B).

3.5. Analysis of grain structure and distribution of AGP

Immunolabelling of sections of developing wheat caryopses with the LM30 monoclonal antibody showed labelling of the starchy endosperm cells and aleurone cells (Fig. 6A–D). In null grain (Fig. 6C and D) labelling within the starchy endosperm was widely distributed around the protein bodies which is consistent with a vacuolar location, but labelling was also present around the plasma membrane. In the transgenic lines (Fig. 6A and B) the labelling of the cell contents was abolished while the labelling of the plasma membrane was still observed. Labelling in the aleurone was observed only in the plasma membrane of both null and transgenic lines.

We therefore conclude that the major form of AGP in starchy endosperm cells is located within the cell, probably in the vacuole, and that the labelling of the plasma membrane resulted from the presence of one or more other protein components (carrying AG glycans including the LM30 glycan epitope) which were not encoded by the *Gsp-1* locus.

Staining of the same cells for the cell wall located β-glucan with Calcofluor showed no co-labelling of the AGP epitope and cell walls in either the control or transgenic lines.

3.6. Effects on grain size, shape and composition

A range of parameters were measured to determine whether the

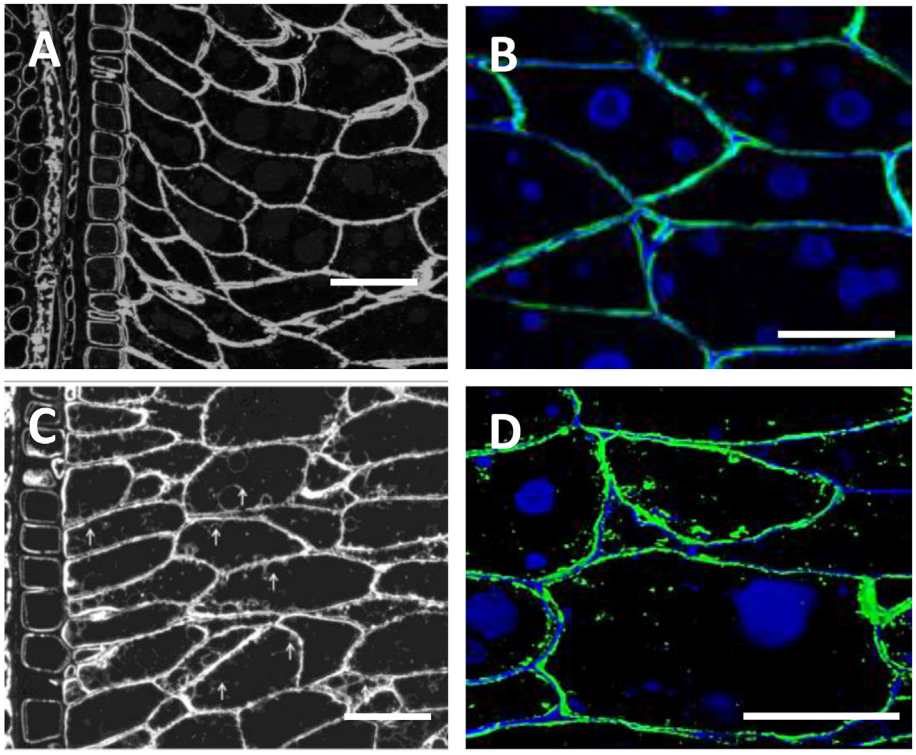

Fig. 6. Immunofluorescence labelling (green) with LM30 antibody of medial transverse sections prepared from 15 DAA *Gsp-1* RNAi wheat grain (A, B) and null segregants (C, D). Micrographs B and D counterstained with Calcofluor White (blue) to demonstrate no co-localisation of the AGP epitope and the cell wall. Bars 100 μm (A), 50 μm (B–D). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

down-regulation of AGP/GSP had wider impacts on grain development and composition. Full details of these are provided in Supplementary material (Supplementary Table S6).

No statistical differences in the 1000 grain weights (TGW) between either type or line were found, with overall means of 39.77 for the null and 39.41 for the transgenic lines. However, the transgenic lines had slightly higher N contents (% dry wt.) than the null lines, the means being 2.55% dry wt. and 2.42% dry wt., respectively (SED = 0.054 on 31 df; LSD (5%) = 0.1106) (Supplementary Table S6). Changes in soluble carbohydrate composition might be expected due to the increased availability of substrate due to the reduction in AGP. We therefore used ^{1}H-NMR to determine sugars (mono- di- and trisaccharides) in polar extracts. This showed statistically significant increases in the concentration of raffinose (a trisaccharide comprising galactose, glucose and fructose) in all transgenic lines compared to the corresponding nulls, and significant increases in galactinol (O-α-D-galactopyranosyl-(1 → 1)-L-myo-inositol) in six lines, sucrose in 5 lines and glucose and maltose in single lines (Supplementary Table 7).

The major arabinose-containing component in wheat grain is arabinoxylan (AX), which is the major cell wall polysaccharide accounting for up to 3% of the dry weight of white flour (Gebruers et al., 2008). Although some differences were observed between the total amounts of AX in the RNAi lines and their corresponding null lines (Supplementary Tables 8a–8c) these were not consistent and there were no significant differences across all pairs of lines. However, an increased level of arabinose substitution was observed in the transgenic lines, which may have resulted from increased availability of arabinose substrate due to suppression of AGP synthesis.

No differences in seed germination were observed between the transgenic and null lines with all lines showing 100% germination after 3 days (Supplementary Table S6).

4. General discussion

Although wheat flour AGP was initially characterised over 40 years ago (Fincher and Stone, 1974) and accounts for about 0.4% of the flour dry weight (Loosveld et al., 1997), nothing is known about its biological role or localisation in the cells of the developing and mature grain. Our demonstration that both AGP and GSP are reduced in amount in the *Gsp-1* RNAi lines provides the first direct evidence that the AGP peptide moiety and GSP are encoded by the same gene, as suggested by Van den Bulck et al. (2005). The post-translational processing (including *O*-glycosylation and proteolysis) of the pre-protein encoded by the transcript are consistent with its trafficking within the endomembrane system into the vacuole, and our data on immunolocalisation are consistent with this. Hence, we conclude that the pre-protein is transported *via* the endoplasmic reticulum and Golgi into the vacuole where AGP and GSP are released by proteolysis. Both then remain within the vacuole, with the GSP associating with the surface of the vacuolar-derived protein bodies.

However, it is clear from the western blotting, ^{1}H-NMR analysis and immunolocalisation that wheat grain contains other forms of arabinogalactan (AG) proteins or peptides detected by the LM30 antibody. This is to be expected as AG proteins are ubiquitous in plants, where they are often involved in cell communication and signalling (Showalter, 2001). The strong labelling of the plasma membrane in the control and transgenic lines is consistent with the previous demonstration in wheat endosperm of transcripts encoding fasciclin-like AG proteins with glycosylphosphoinositol (GPI) anchors to the plasma membrane (Faik et al., 2006). However, although some AG proteins are located in the cell walls (Showalter, 2001) we found no evidence for labelling of the wheat starchy endosperm cell walls with LM30.

GSP was initially identified as a determinant of grain texture (Jolly et al., 1996). However, this has since been disputed, with Tranquilli et al. (2002) showing that variation in *Gsp-1* (resulting from the effects of differences in gene dosage, isogenic substitution or allelic variation) did not have any effect on grain texture. However, we also know that allelic variation in Pin a and Pin b determines only 60–80% of the variation in grain texture in bread wheat, while variant forms of Pin b may account for another 10% (by 5–7 SKCS units) (Wilkinson et al., 2008). We show here that down-regulation of *Gsp-1* also results in a small, but consistent, reduction in grain hardness, although the mechanism, or even whether it relates to effects on AGP, GSP or both components, is not known. Finally, down-regulation of *Gsp-1* also results in a small decrease in the viscosity of aqueous extracts. Viscosity is an important quality trait, affecting processing, health benefits, feed conversion in livestock and the efficiency of brewing, distilling and biofuel production (Pettersson and Åman, 1989; Courtin and Delcour, 2002; Topping, 2007).

Acknowledgements

Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK. We are grateful to Bioimaging at Rothamsted Research for help with the preparation of sections for immunofluorescence slides and to Mr David Steele and Mr. Andrew Riche for carrying out the nitrogen percentage measurements and calculations. We are grateful to Dr. Didier Marion (INRA, Nantes) for providing the GSP antibody.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jcs.2017.02.006.

References

Baker, J.M., Hawkins, N.D., Ward, J.L., Lovegrove, A., Napier, J.A., Shewry, P.R., Beale, M.H., 2006. A metabolomic study of substantial equivalence of field-grown genetically modified wheat. Plant Biotechnol. J. 4, 381–392.

Bazin, H., 1982. Production of rat monoclonal antibodies with the LOU rat non-secreting IRF983F myeloma cell line. Protoc. Bio Fluids 29, 615–618.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P., Chalhoub, B., 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). Plant Cell 17, 1033–1045.

Charmet, G., Masood-Quraishi, U., Ravel, C., Romeuf, I., Balfourier, F., Perretant, M.R., Joseph, J.L., Rakszegi, M., Guillon, F., Sado, P.E., Bedo, Z., Saulnier, L., 2009. Genetics of dietary fibre in bread wheat. Euphytica 170, 155–168.

Courtin, C.M., Delcour, J.A.J., 2002. Arabinoxylans and endoxylanases in wheat flour bread-making. J. Cereal Sci. 35, 225–243.

Elmorjani, K., Geneix, N., Dagalarrondo, M., Branlard, G., Didier, M., 2013. Wheat grain softness protein (*Gsp-1*) is a puroindoline-like protein that displays a specific post-translational maturation and does not interact with lipids. J. Cereal Sci. 58, 117–122.

Faik, A., Abouzauhir, J., Sarham, F., 2006. Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (*Triticum* aestivum) and rice (*Oryza* sativa): identification and bioinformatics analyses. Mol. Genet. Genomics 276, 478–494.

Fincher, G.B., Stone, B.A., 1974. Water-soluble arabinogalactan-peptide from wheat endosperm. Aust. J. Biol. Sci. 27, 117–132.

Gebruers, K., Dornez, E., Boros, D., Fraś, A., Dynkowska, W., Bedo, Z., Rakszegi, M., Delcour, J.A., Courtin, C.M., 2008. Variation in the content of dietary fiber and components thereof in wheats in the HEALTHGRAIN Diversity Screen. J. Agric. Food Chem. 56 (21), 9740–9749.

Jolly, C.J., Glenn, G.M., Rahman, S., 1996. *GSP-1* genes are linked to the grain hardness locus (*Ha*) on wheat chromosome 5D. Proc. Natl. Acad. Sci. U. S. A. 93, 2408–2413.

Loosveld, A.-M.A., Grobet, P.J., Delcour, J.A., 1997. Contents and structural features of water-extractable arabinogalactan in wheat flour fractions. J. Agric. Food Chem. 45, 1998–2002.

Mead, R., Curnow, R.N., Hasted, A.M., 2003. Statistical Methods in Agriculture and Experimental Biology. Chapman and Hall, London, UK, p. 472.
Morris, C.F., Rose, S.P., 1996. Wheat. In: Henry, R.J., Kettlewell, P.S. (Eds.), Cereal Grain Quality. Chapman Hall, London), pp. 3–54.
Nemeth, C., Freeman, J., Jones, H.D., Sparks, C., Pellny, T.K., Wilkinson, M.D., Dunwell, J., Andersson, A.A.M., Aman, P., Guillon, F., Saulnier, L., Mitchell, R.A.C., Shewry, P.R., 2010. Downregulation of the CSLF6 gene results in decreased (1,3;1,4)-β-D-glucan in endosperm of wheat. Plant Physiol. 152, 1209–1218.
Pettersson, D., Åman, P., 1989. Enzyme supplementation of a poultry diet containing rye and wheat. Br. J. Nutr. 62, 139–149.
Rahman, S., Jolly, C.J., Skerritt, J., Wallosheck, A., 1994. Cloning of a wheat 15 kDa grain softness protein (GSP). GSP is a mixture of different puroindoline-like polypeptides. Eur. J. Biochem. 223, 917–925.
Rieu, I., Powers, S.J., 2008. Real time quantitative RT-PCR: design, calculations and statistics. Plant Cell 21, 1031–1033.
Saulnier, L., Peneau, N., Thibault, J.F., 1995. Variability in grain extract viscosity and water-soluble arabinoxylan content in wheat. Journal of Cereal Science 22, 259–264.
Shewry, P.R., Halford, N.G., 2002. Cereal seed storage proteins: structures, properties and role in grain utilization. J. Exp. Bot. 53, 947–958.
Showalter, A.M., 2001. Arabinogalactan-proteins: structure, expression and function. Cell. Mol. Life Sci. 58, 1399–1417.
Sparks, C.A., Jones, H.D., 2009. Biolistics transformation of wheat. In: Jones, H.D., Shewry, P.R. (Eds.), Transgenic Wheat, Barley and Oats. Humana Press, New York, NY, pp. 71–92.
Topping, D., 2007. Cereal complex carbohydrates and their contribution to human health. J. Cereal Sci. 46, 220–229.
Tosi, P., Parker, M., Gritsch, C.S., Carzaniga, R., Martin, B., Shewry, P.R., 2009. Trafficking of storage proteins in developing grain of wheat. J. Exp. Bot. 60 (3), 979–991.
Tranquilli, G., Heaton, J., Chicaiza, O., Dubcovsky, J., 2002. Substitutions and deletions of genes related to grain hardness in wheat and their effect on grain texture. Crop Sci. 42, 1812–1817.
Tryfona, T., Liang, H.-C., Kotake, T., Kaneko, S., Marsh, J., Ichinose, H., Lovegrove, A., Tsumuraya, Y., Shewry, P.R., Stephens, E., Dupree, P., 2010. Carbohydrate structural analysis of wheat flour arabinogalactan protein. Carbohydr. Res. 345, 2648–2656.
Turner, A.S., Bradburne, R.P., Fish, L., Snape, J.W., 2004. New quantitative trait loci influencing grain texture and protein content in bread wheat. J. Cereal Sci. 40, 51–60.
Van den Bulck, K., Swennen, K., Loosveld, A.-M.A., Courtin, C.M., Brijs, K., Proost, P., Van Damme, J., Van Campenhout, S., Mort, A., Delcour, J.A., 2005. Isolation of cereal arabinogalactan-peptides and structural comparison of their carbohydrate and peptide moieties. J. Cereal Sci. 41, 59–67.
Wilkinson, M., Wan, Y., Tosi, P., Leverington, M., Snape, J., Mitchell, R.A.C., Shewry, P.R., 2008. Identification and genetic mapping of variant forms of puroindoline b expressed in developing wheat grain. J. Cereal Sci. 48, 722–728.
Wilkinson, M.D., Castells-Brooke, N., Shewry, P.R., 2013. Diversity of sequences encoded by the *Gsp-1* genes in wheat and other grass species. J. Cereal Sci. 57, 1–9.
Willats, W.G.T., Marcus, S.E., Knox, J.P., 1998. Generation of monoclonal antibody specific to (1→5)-Alpha-L-arabinan. Carbohydr. Res. 308, 149–152.